

1 判断木構成手順の支援

ある属性を選んだ際に学習データセットがどのように分類されるのかを計算するプログラムである「decision.c」を示します。decision.c プログラムを用いると、ある学習データセットの分類にはどの属性を用いるべきかを決定することができます。すなわち、各属性を用いた場合の分類結果を比べることで、どの属性を用いて分類すべきかを定めることができます。ある属性を決定した後、分類後のデータについて再び decision.c プログラムを適用することで、次にどの属性を用いて分類すべきかを定めることができます。これを繰り返すことで、判断木を構成します。

はじめての機械学習 小高知宏著 106 頁より

2 判断木作成支援プログラム

decision.c

```
/*
  decision.c
  判断木の構成を支援するプログラム
  入力は標準入力から与え、出力は標準出力に出力します。
  使い方
  ./decision <(入力ファイル)> <(出力ファイル)>
  入力ファイルは、属性とカテゴリを 0/1 で記述します。
  出力ファイルは、分類結果を表示します。
*/

#include <stdio.h>
#include <stdlib.h>
#include <string.h>

#define TEXTLENGTH 256
#define DATANO 256
#define ATTNO 4
#define YES 1
#define NO 0

int readdata(int dataset [DATANO][ATTNO + 1]);
void evalp(int dataset [DATANO][ATTNO + 1], int no, int i, int yorn);

int main(int argc, char *argv[])
{
  int no;
  int i;
  int j;
  int dataset [DATANO][ATTNO + 1] = {0};
```

```

no = readdata(dataset);

for(i = 0; i < ATTNO; ++i){
    printf("\np%d\n", i + 1);
    printf("Yes ");
    evalp(dataset, no, i, YES);

    printf("No ");
    evalp(dataset, no, i, NO);
}

return 0;
}

int readdata(int dataset [DATANO] [ATTNO + 1])
{
    int no = 0;
    int att = 0;
    int data;

    while(scanf("%d", &data) != EOF){

        dataset[no][att] = data / 10000;
        ++att;
        dataset[no][att] = (data % 10000)/ 1000;
        ++att;
        dataset[no][att] = (data % 1000)/ 100;
        ++att;
        dataset[no][att] = (data % 100)/ 10;
        ++att;
        dataset[no][att] = (data % 10);

        ++no;
        att = 0;
    }

    return no;
}

void evalp(int dataset [DATANO] [ATTNO + 1], int no, int i, int yorn)
{
    int j;
    int total;
    int point;

    j = 0;
    total = 0;
    point = 0;

```

```

while(j < no){
    if(dataset[j][i] == yorn){
        ++total;

        if(dataset[j][i] == dataset[j][ATTNO])
            ++point;
    }
    ++j;
}
printf("%d/%d %lf\n", point, total, (double)point/total);
}

```

3 「decision.c」が受け取るデータ ex.txt

```

11101
10101
01101
11101
00100
11000
11110
11000
11010
00110

```

4 「decision」の操作方法

「decision」の操作方法の例

```
./decision < ex.txt
```

上記の「./」は、Linuxでのプログラムを実行する時の例です。
 ウィンドウズでは、decision < ex.txt「Enter」と入力してください。

5 「decision」の実行結果の例

```

p1 Yes 3/7 0.428571 No 2/3 0.666667
p2 Yes 3/7 0.428571 No 2/3 0.666667

```

p3 Yes 4/7 0.571429 No 3/3 1.000000

p4 Yes 0/3 0.000000 No 3/7 0.428571